

Partial Translation of JP 1990-3520

Publication Date: January 23, 1990

Application No.: 1982-65142

Filing Date: April 16, 1982

Applicant: MITSUBISHI ELECTRIC CORP

Inventor: Kenji SHIMA

Inventor: Masahiro HIBINO

Lines 4-18, column 11, page 6

FIG. 8 is a specific block diagram of the audio recognition circuit 29 included in FIG. 6. The audio recognition circuit 29 shown in FIG. 8 uses a conventionally known channel filter, configuration and operation of which are briefly described below. To the input blocking switch 291 is given an audio input signal from the directional microphone 19 as an input signal, from the filter 28 as shown in FIG. 6. To the input blocking switch 291 is input an audio noise judgment signal from the comparator 26 as shown in FIG. 6, as a control signal. Then, when the audio noise judgment signal represents sound, the input blocking switch 291 will be closed and recognition operation enabled. When it is judged that it does not represent sound, i.e., it represents noise, it will be opened and recognition operation stopped.

⑫ 特 許 公 報 (B 2)

平2-3520

⑮ Int. Cl. ⁵	識別記号	庁内整理番号	⑭ 公告 平成2年(1990)1月23日
G 10 L 3/00	3 0 1 A	8842-5D	
B 60 R 16/02		7443-3D	
G 10 L 3/00	3 0 1 Z	8842-5D	
H 04 R 3/00	3 2 0	8524-5D	
// G 10 L 9/08	3 0 1 A	8842-5D	

発明の数 1 (全19頁)

⑭ 発明の名称 音声識別装置

⑮ 特 願 昭57-65142

⑯ 公 開 昭58-181099

⑮ 出 願 昭57(1982)4月16日

⑯ 昭58(1983)10月22日

⑰ 発 明 者 嶋 憲 司 兵庫県尼崎市南清水字中野80番地 三菱電機株式会社応用機器研究所内

⑰ 発 明 者 日 比 野 昌 弘 兵庫県尼崎市南清水字中野80番地 三菱電機株式会社応用機器研究所内

⑱ 出 願 人 三菱電機株式会社 東京都千代田区丸の内2丁目2番3号

⑲ 代 理 人 弁理士 大 岩 増 雄

審 査 官 中 村 和 男

⑳ 参 考 文 献 特開 昭56-119200 (JP, A)

1

2

⑳ 特許請求の範囲

1 発音者の発音した音声を識別するための音声識別装置であつて、

それぞれが前記発音者の近傍に所定の間隔を有して配置され、入力された音声を電気信号に変換する少なくとも第1および第2の音声入力手段を含む音声電気信号変換手段、

前記第1の音声入力手段から出力される音声入力信号から予め定める周波数帯域成分の音声信号を抽出する第1のフィルタ手段、

前記第2の音声入力手段から出力される音声入力信号から前記第1のフィルタ手段と同一の周波数帯域成分の音声信号を抽出する第2のフィルタ手段、

前記第1および第2のフィルタ手段から抽出されたそれぞれの音声信号の相関を演算する相関演算手段、および

前記相関演算手段からの演算結果に基づいて、前記音声電気信号変換手段から出力される音声信号が前記発音者の発音した音声に基づくものであるかあるいは騒音に基づくものであるかを判別する音声判別手段を備えた、音声識別装置。

2 前記音声判別手段は、

予め定める音声対騒音比としての値を記憶する記憶手段、

前記記憶手段に記憶している音声対騒音比と前記相関演算手段出力とを比較し、音声対騒音比の5 大小を表わす信号を出力する比較手段、および前記比較手段からの音声対騒音比の大小を表わす信号に基づいて、前記音声電気信号変換手段から入力された音声信号の音声認識を行なうかあるいは音声認識を不能にする音声認識手段を含む、10 特許請求の範囲第1項記載の音声識別装置。

3 前記相関演算手段は、前記第1および第2のフィルタ手段から出力された音声入力信号のそれぞれの時間差に対応した相関係数を演算する相関係数演算手段を含み、

15 前記記憶手段は予め定める相関係数を音声対騒音比対応信号として記憶し、

前記比較手段は前記相関係数演算手段出力と前記記憶手段に記憶している相関係数とを比較して発音者の発音した音声と騒音との比に応じた音声騒音判定信号を出力するようにした、特許請求の20 範囲第2項記載の音声識別装置。

4 前記音声判別手段は、前記相関係数演算手段出力から相関係数の最大

値を有する時間を検出する遅れ時間検出手段と、

前記第1および第2の音声入力手段出力の音声入力信号のそれぞれの時間差を中央値とし、それから所定の時間の範囲の値を記憶する時間範囲記憶手段と、

前記遅れ時間検出手段によつて検出された遅れ時間と前記時間範囲記憶手段の内容とを比較し、遅れ時間が前記時間範囲記憶手段の内容に基づく範囲内に含まれていることに応じて音声判定信号を出力し、前記範囲外であれば騒音判定信号を出力する第2の比較手段と、

前記第2の比較手段からの音声判定信号に応じて前記音声電気信号変換手段から入力された音声入力信号の認識を能動化し、前記騒音判定信号に応じて前記認識を不能化する第2の音声認識手段とを含む、特許請求の範囲第1項記載の音声識別装置。

5 前記第1または第2の音声認識手段は、

前記音声電気信号変換手段出力に基づいて音声パラメータを抽出する音声パラメータ抽出手段と、

予め定める音声パラメータを記憶する音声パラメータ記憶手段と、

前記音声パラメータ抽出手段からの音声パラメータと前記音声パラメータ記憶手段に記憶している音声パラメータとの類似度を演算する類似度演算手段と、

前記類似度演算手段による類似度の演算結果に基づいていずれの音声であるかを判定するキーワード判定手段とを含み、

前記類似度演算手段は、前記相関係数演算手段出力の相関係数に基づく音声対騒音比対応信号に応じた重み関数を考慮して類似度演算を行なうようにした、特許請求の範囲第2項または第4項に記載の音声識別装置。

6 前記重み関数は2値信号である、特許請求の範囲第5項記載の音声識別装置。

7 前記音声電気信号変換手段は、前記音声判別手段に音声信号を与える第3の音声入力手段を含む、特許請求の範囲第1項記載の音声識別装置。

8 前記第3の音声入力手段は指向性マイクロホンを含む、特許請求の範囲第7項記載の音声識別装置。

9 前記第1および第2の音声入力手段はそれぞ

れ前記発音者の口許から等距離の位置に配置される、特許請求の範囲第1項記載の音声識別装置。

発明の詳細な説明

この発明は音声識別装置に関し、特に、音声を入力し、その音声の内容に基づいて各種被制御機器を制御するために用いられるような音声識別装置の改良に関する。

たとえば、自動車などにおいて運転者がラジオなどの車載機器を操作する場合には、一般に手動的に行なうことが多い。しかし、運転者が車を運転中にラジオなどを操作するためには、しばしば視線を前方から機器の操作部に移す必要があり、このことが脇見運転の原因になり、非常に危険を判う。また、最近では運転者用の車載用計算機が実用に供されているが、このような車載用計算機を、運転しながら操作することは不可能に近いものであった。この他にも、窓の開閉やその他の車載機器を操作する場合においても同様の問題点を生じていた。

そこで、運転者が手動的な操作によることなく、ラジオの電源投入や窓の開閉などをすることができれば便利である。このためには、運転者が「ラジオ」と発音したとき、この音声を識別してその識別結果に基づいてラジオを制御するような音声入力制御装置が考えられる。このような音声入力制御装置では、運転者の近傍にマイクロホンを設けておき、このマイクロホンに入力された音声を識別する。ところが、自動車の窓が開かれていると、外部からの騒音がマイクロホンに混入してしまい、運転者の発音と騒音との識別が極めて困難になり、車載機器などを誤動作させる原因となっていた。

それゆえに、この発明の主たる目的は、外部からの騒音と人間の発音した音声とを容易に識別し得る音声識別装置を提供することである。

この発明を要約すれば、発音者の近傍に所定の間隔を隔てて第1および第2の入力手段を配置しておき、それぞれの音声入力手段から出力された音声入力信号から予め定める周波数帯域成分の音声信号を抽出し、それぞれ音声信号の相関を演算し、その演算結果に基づいて音声電気信号変換手段から出力された音声信号が発音者の発音した音声に基づくものであるかあるいは騒音に基づくものであるかを判別するように構成したものであ

る。

この発明の上述の目的およびその他の目的と特徴は以下に図面を参照して行なう詳細な説明から一層明らかとなろう。

第1図はこの発明の一実施例が適用された自動車を横方向から見た概略図であり、第2図は同じく運転者の後方向から見た概略図であり、第3図は同じく自動車の車内の概略図である。

まず、第1図ないし第3図を参照して、この発明の概要について説明する。運転者11は車体12に設けられている座席13に着席し、シフトレバー14およびステアリングホイール15を操作して運転する。計器盤16にはたとえばカーラジオ7が取付けられている。また、天井20には、運転者11の正面方向に直交して等距離の位置に所定の間隔を有して第1および第2の音声入力手段としてのマイクロホン17と18とが取付けられている。さらに、これらのマイクロホン17と18との間には指向性マイクロホン19が設けられている。

運転者11がたとえばラジオ7の電源を投入しようとする場合は、予め定められている制御用の音声（以下、キーワードと称する）を発音する。この発音は空气中を伝播し、マイクロホン17と18と19とに入る。このとき、マイクロホン17と18が運転者11から等距離の位置に設けられているので、音声はこれらのマイクロホン17と18とに同時に伝播されることになる。マイクロホン17と18とに入力された音声は電気信号に変換されて音声信号となり、この2つの音声信号の相互関係数が求められる。騒音が入力されていない場合には、相関係数は1に近い値が得られ、音声信号は騒音に比べて十分に大きいことが相関係数の値から示される。逆に、大きな騒音が入力された場合には相関係数は0に近い値が得られ、音声信号は騒音に比べて十分に小さいことが示される。そして、この相互相関係数の値に基づいて、指向性マイクロホン19から出力される音声信号を認識するか否かを判別する。そして、相互相関係数が1に近い値であれば、指向性マイクロホン19から出力される音声信号と予め記憶されているキーワードの音声パラメータとを比較し、いずれのキーワードであるかを判別する。そして、この判別結果に基づいてたとえばラジオ7

などの各種車載機器を制御する。

なお、上述の説明では、音声認識用の音声信号を指向性マイクロホン19から出力するようにしたが、特に指向性マイクロホン19を設けることなく、マイクロホン17または18のいずれか一方からの音声信号を認識するようにしてもよい。しかしながら、運転者11の発音した音声を分析して認識を行なうための音声信号のS/N比としては十分高いものが好ましく、周囲雑音を除去できる高指向性マイクロホンをを用いるのが望ましい。

第4図は音声信号と騒音との関係を説明するための図であり、第5A図および第5B図は相互相関係数の計算結果を示す図である。

次に、2つのマイクロホン17と18とから出力される音声信号の相互相関について説明する。第4図において、騒音発生源100として、この場合は一例として自動車騒音について考えてみる。運転者11はマイクロホン17と18とから等距離に口許がくる位置に座っており、運転者11の前方には自動車のフロントガラス（図示せず）、後方にもリアウィンドウ（図示せず）が存在するものとし、騒音発生源100は運転者11の側方にあるものとする。

騒音発生源100から騒音がマイクロホン17と18とに伝播する遅延時間差を τ_0 とし、その振幅差はマイクロホン17と18との間隔dに比べて、マイクロホン17と18とから騒音源100の距離が十分に大きいものとみなして考慮せずに振幅が同一であるとする。また、運転者11により発生された音声信号 S_a 、 S_b はマイクロホン17と18とに同一の振幅と時間で伝播するものとする。このとき、マイクロホン17と18との時間信号 $a(t)$ と $b(t)$ は、音声信号を $s(t)$ 、騒音発生源100からの騒音 $n(t)$ とすると、

$$a(t) = s(t) + n(t) \quad \dots(1)$$

$$b(t) = s(t) + n(t + \tau_0) \quad \dots(2)$$

と表わすことができる。ここで、相互相関関数 $\phi_{ab}(\tau)$ および正規化相互相関関数 $\Phi(\tau)$ はそれぞれ、

$$\phi_{ab}(\tau) = \lim_{T \rightarrow \infty} 1/2T \int_{-T}^T a(t) b(t + \tau) dt \quad \dots(3)$$

$$\Phi(\tau) = \phi_{ab}(\tau) / \sqrt{\phi_{aa}(0) \phi_{bb}(0)} \quad \dots(4)$$

である。但し、 $\phi_{aa}(0)$ 、 $\phi_{bb}(0)$ は、 $a(t)$ 、

7

8

$b(t)$ の自己相関関数 $\phi_{aa}(\tau)$ 、 $\phi_{bb}(\tau)$ のそれぞれ $\tau = 0$ のときの値であつて、パワーを表わしており、

$$\phi_{aa}(0) = \lim_{T \rightarrow \infty} 1/2T \int_{-T}^T a^2(t) dt \quad \dots(5)$$

$$\phi_{bb}(0) = \lim_{T \rightarrow \infty} 1/2T \int_{-T}^T b^2(t) dt \quad \dots(6)$$

で表現できる。ここで、

$$s(t) = S \sin \omega_s t$$

$$n(t) = N \sin \omega_n t$$

とする。但し、 ω_s 、 ω_n は $s(t)$ 、 $n(t)$ の角周波数であり、 S 、 N は $s(t)$ 、 $n(t)$ の振幅を示している。そして、 $\omega_s = 2\pi f_s$ 、 $\omega_n = 2\pi f_n$ であるとすると、正規化相互相関関数 $\Phi(\tau)$ は、*

$$\Phi(\tau) = \frac{\sum_{l=1}^L S^2 \cos \omega_s \tau + \sum_{l=1}^M N^2 \cos \omega_n (\tau + \tau_0)}{\sum_{l=1}^L S^2 + \sum_{l=1}^M N^2} \quad \dots(11)$$

であり、また $\tau = 0$ とすれば、

$$\Phi(0) = \frac{\sum_{l=1}^L S^2 + \sum_{l=1}^M N^2 \cos \omega_n \tau_0}{\sum_{l=1}^L S^2 + \sum_{l=1}^M N^2} \quad \dots(12)$$

である。

前述の第(8)式において、 $\tau_0 = 333 \mu s$ つまりマイクロホン 17 と 18 との間隔 d が約 $1.13 cm$ における計算結果を第 5 A 図に示している。この第 5 A 図から明らかなように、側方から入ってくる騒音が大きくなりかつ S/N が小さくなるに従い正規化相関関数の小さくなることがわかる。但し、正弦波の場合には、 $1/\tau_0$ の周波数の整数倍に近い周波数の騒音の場合には、第 5 A 図に示したようにはならないことは明らかであるので、実際にはそのような成分を小さくした状態で求める必要がある。

また、前述の第(12)式において、 $\tau_0 = 333 \mu s$ 、 $d = 1.13 cm$ における計算結果を第 5 B 図に示している。この第 5 B 図では、騒音 $n(t)$ を

$$\begin{aligned} n(t) = & 0.4 \sin(2\pi \cdot 100) t \\ & + 1.0 \sin(2\pi \cdot 400) t \\ & + 1.0 \sin(2\pi \cdot 1000) t \\ & + 1.0 \sin(2\pi \cdot 3100) t \end{aligned} \quad \dots(13)$$

$$\Phi(\tau) = \{S^2 \cos \omega_s \tau + N^2 \cos \omega_n (\tau + \tau_0)\} / (S^2 + N^2) \quad \dots(7)$$

であり、また $\tau = 0$ とすれば、

$$\Phi(0) = (S^2 + N^2 \cos \omega_n \tau_0) / (S^2 + N^2) \quad \dots(8)$$

である。

さらに、 $s(t)$ 、 $n(t)$ がそれぞれ L 個、 M 個の正弦波から成立っている場合には、

$$s(t) = \sum_{l=1}^L S_l \sin \omega_{sl} t \quad \dots(9)$$

$$n(t) = \sum_{l=1}^M N_l \sin \omega_{nl} t \quad \dots(10)$$

となるので、

とし、100Hz、400Hz、1kHz、3.1kHz の 4 つの周波数成分が 0.4、1、1、1 の比率で含まれているものとした。この場合には、 $\tau_0 = 333 \mu s$ の周期に近い 3.1kHz の成分や低い周波数成分を含んでいるので、 $\Phi(0)$ はさほど小さくはならないが、それでも S/N が小さくなると、 $\Phi(0)$ が小さくなる傾向が見られる。

この発明は 2 つのマイクロホン 17 と 18 とから出力される音声信号の相関関係を用いて音声信号の S/N 比を推定し、その値によつて認識結果の有効あるいは無効を決定したり、類似度計算に重み付けたりすることによつて、騒音による誤認識や誤判定を防止できるような音声識別装置を構成できる。以下に、この発明の実施例について詳細に説明する。

第 6 図はこの発明の一実施例の概略ブロック図である。前述の第 2 図および第 3 図に示したマイクロホン 17 からの音声信号は増幅器 21 で増幅され、フィルタ 22 によつて識別したい周波数帯

域（通常は音声周波数帯域）のみが抽出され、相互相関係数計算機 25 に与えられる。他方のマイクロホン 18 からの音声信号は同様に増幅器 23 で増幅され、フィルタ 24 によつて識別したい周波数帯域成分のみが抽出されて相互相関係数計算機 25 に与えられる。この相互相関係数計算機 25 は入力された 2 つの音声信号から相互相関係数 $\Phi(O)$ を求めるものである。この相互相関係数計算機によつて計算された相互相関係数 $\Phi(O)$ の値によつて S/N を推定することができる。このために、相互相関係数計算機 25 で計算された相互相関係数 $\Phi(O)$ が比較器 26 に与えられる。比較器 26 には予め定める相関係数を記憶するための相関係数しきい値レジスタ 27 が接続される。そして、比較器 26 は相互相関係数計算機 25 で計算された相互相関係数 $\Phi(O)$ の値と相関係数しきい値レジスタ 27 に記憶されている相関係数しきい値と比較し、指向性マイクロホン 19 から出力される音声信号を認識するために騒音が小さい否かの判定を行なう。そして、その判定

信号を音声認識認識回路 29 に与える。

一方、指向性マイクロホン 19 は 2 つのマイクロホン 191 と 192 とを含み、それぞれから出力された音声信号は増幅器 193 および 194 で増幅され、加算器 195 で加算されてフィルタ 28 に与えられる。このフィルタ 28 は音声信号として不必要な低周波域や高周波領域の信号成分を除去するためのものである。フィルタ 28 の出力信号は音声認識回路 29 に与えられる。音声認識回路 29 は後述の第 8 図で詳細に説明するが、入力された騒音の混入した音声信号から特徴パラメータを抽出し、メモリ 38 に予め登録されている音声パラメータとの類似度を求め、その類似度が一致していると判断するのに十分でない場合には、入力された音声信号を判定できないことを出力し、最も類似度が大きくかつ類似度が十分に大きい場合にはそのデータであると判定して出力するものである。

音声認識回路 29 には比較器 26 から音声の騒音判定信号が与えられており、この信号が騒音であることを示している場合には、音声認識回路 29 は騒音または判定不能を表わす信号を出力することによつて、騒音が大きいときに各種機器、この実施例ではラジオ 7 が誤動作しないように構成

される。

第 7 A 図は第 6 図に示す相互相関係数計算機 25 の一例を示す概略ブロック図である。第 7 A 図において、 $a(t)$ はマイクロホン 17 から出力される音声信号であり、 $b(t)$ はマイクロホン 18 から出力される音声信号である。そして、乗算器 251 と積分器 254 と対数変換器 257 は前述の第(5)式を演算するものである。すなわち、乗算器 251 は a を 2 乗し、これを積分器 254 で積分し、対数変換器 257 によつて対数化し、対数化された $\phi 3a(O)$ を演算する。また、乗算器 252 乗算器 256 と対数変換器 259 は前述の第(6)式の演算を行なうものである。すなわち、乗算器 252 は b を 2 乗し、積分器 256 で積分し、対数変換器 259 は先の対数変換器 257 と同様に対数化された $\phi bb(O)$ を求める。対数変換器 257 と 259 との出力は加算器 260 で加算され、1/2減衰器 261 によつてその平方根に対応した処理が行なわれる。したがつて、1/2減衰器 261 の出力には、前述の第(4)式の分母に対応した結果が得られることになる。

一方、乗算器 253 と積分器 255 と対数変換器 258 は前述の第(3)式の $\tau = 0$ のときの演算を行なうものである。すなわち、乗算器 253 は a と b とを乗算し、それを積分器 255 が積分し、対数変換器 258 は第(3)式の $\phi ab(O)$ の対数化された値を出力する。この出力は加算器 262 に与えられ、1/2減衰器 261 の出力と減算され、指数変換器 263 によつて指数化してもとに戻され、 $\Phi(O)$ が出力される。すなわち、この第 7 A 図に示す相互相関係数計算機 25 は、マイクロホン 17 および 18 から出力される音声信号に基づいて、騒音が小さい場合には $\Phi(O) = 1$ を出力し、騒音が大きい場合には $\Phi(O) = 1$ を出力する。

第 7 B 図は相互相関係数計算機 25 の他の例を示す概略ブロック図である。第 7 B 図において、2乗平均値回路 (RMS) 266 は時間成分 a の 2 乗平均を開根するものであり、 $\sqrt{a^2}$ を演算する。加算器 264 は a と b とを加算し、RMS 267 によつてその 2 乗平均値 $\sqrt{(a+b)^2}$ が演算される。加算器 265 は a と $-b$ とを加算するものであり、RMS 268 はその結果の 2 乗平均値 $\sqrt{(a-b)^2}$ を演算する。RMS 268 は b の 2 乗平均値

$\sqrt{B^2}$ を演算する。各RMS 2 6 6ないし2 6 9の演算結果は演算回路2 6 9によつて $A^2-B^2/4CD$ を演算し、 $\Phi(O)$ を求める。

第8図は第6図に含まれる音声認識回路2 9の具体的なブロック図である。この第8図に示す音声認識回路2 9は従来から知られたチャネルフィルタを用いたものであつて、以下その構成と動作について簡単に説明する。入力遮断スイッチ2 9 1には入力信号として第6図に示すフィルタ2 8から指向性マイクロホン1 9からの音声入力信号が与えられる。また、入力遮断スイッチ2 9 1には制御信号として第6図に示す比較器2 6から音声騒音判定信号が入力される。そして、入力遮断スイッチ2 9 1は音声騒音判定信号が音声を表わす信号であるとき閉じられて認識動作を行ない得る状態となり、音声でないつまり騒音であると判定されているときには開いて認識動作を停止した状態となる。このような機能を有する入力遮断スイッチ2 9 1を介して音声信号がブリエンプアシス回路2 9 2に与えられる。ブリエンプアシス回路2 9 2は入力された音声信号のうちの300ないし5kHzの帯域成分のみを通過させる6dB/oct程度の特性を有するフィルタである。

ブリエンプアシス回路2 9 2を通過した音声信号は、バンドパスフィルタ(BPF) 3 0 1ないし3 0 8に与えられる。これらのバンドパスフィルタ3 0 1ないし3 0 8はそれぞれ予め定められた帯域成分の音声信号のみを通過させて、後段の整流回路3 1 1ないし3 1 8に与える。この整流回路3 1 1ないし3 1 8はそれぞれバンドパスフィルタ3 0 1ないし3 0 8の出力信号を整流して直流電圧に変換する。整流された直流電圧はローパスフィルタ3 2 1ないし3 2 8に与えられる。これらのローパスフィルタ3 2 1ないし3 2 8は、一種の積分機能を有したものである。さらに、入力された信号に重み付けをして加算し、その信号をトリガ回路3 4に与える。トリガ回路3 4は入力された信号の変化に応じて音声の始端と終端を検出し、CPU 3 5に対して音声期間であることを表わす信号を与える。また、ローパスフィルタ3 2 1ないし3 2 8のそれぞれ出力信号はマルチプレクサ3 3に与えられる。マルチプレクサ3 3には、CPU 3 5から入出力インタフェース3 6を介してローパスフィルタ3 2 1ないし3

2 8のそれぞれの出力信号を切替るための切替信号が与えられる。そして、マルチプレクサ3 3はその切替信号に応答して、いずれかのローパスフィルタの出力をA-D変換器3 7に与える。A-D変換器3 7は、入力されたアナログ電圧をデジタル値に変換するものである。このデジタル値に変換された信号は、入出力インタフェース3 6を介してCPU 3 5に与えられる。メモリ3 8は音声認識に必要なプログラムを記憶するプログラムメモリ3 8 1と、音声認識に必要なデータを記憶する音声パラメータメモリ3 8 2を含む。音声パラメータメモリ3 8 2は予め登録された分析パラメータを記憶する登録部として記憶領域3 8 3と、入出力インタフェース3 6から出力された音声パラメータを記憶する入力部としての記憶領域3 8 4と、200msec程度の信号の分析パラメータを記憶するためのバッファ部としての記憶領域3 8 5を含む。200msec程度の信号の分析パラメータを記憶するのは、音声認識をする上において音声期間としてはトリガ回路3 4によつて音声期間であると判定された時点以前も極めて重要であるので、類似度を計算する際にはバッファ部3 8 5の内容から始点部を溯つて検出するためである。

次に、動作について説明する。入力遮断スイッチ2 9 1は第6図に示す比較器2 6から騒音であることを表わす判定信号が入力されるとその接点を開き、音声であることを表わす判定信号が入力されればその接点を閉じる。接点が閉じられると、音声入力信号はブリエンプアシス回路2 9 2、バンドパスフィルタ3 0 1ないし3 0 8、整流回路3 1 1ないし3 1 8およびローパスフィルタ3 2 1ないし3 2 8に順次与えられる。マルチプレクサ3 3はCPU 3 5から入出力インタフェース3 6を介して出力される切替信号に応答して、各ローパスフィルタ3 2 1ないし3 2 8のそれぞれの出力を順次選択してA-D変換器3 7に与える。A-D変換器3 7は入力された信号をデジタル値に変換し、入出力インタフェース3 6を介して音声パラメータメモリ3 8 2の入力部3 8 4に記憶させる。一方、トリガ回路3 4はローパスフィルタ3 2 1ないし3 2 8の信号によつて、音声期間の終了を判断し、CPU 3 5に割込信号を与える。応じて、CPU 3 5は入力部3 8

13

4に記憶した音声パラメータと登録部383に予め記憶している音声パラメータとの内容を比較し、ある値以上の類似度が得られないときには判定不能の出力を行ない、ある値以上の類似度が得られたときには最も類似度の高いキーワードを音声認識データとして入出力インタフェース36經由で出力する。なお、類似度の計算のために用いる音声期間としては、トリガ回路34によつて音声期間であると判定された時点以前も極めて重要であるので、CPU35はバッファ部385に記憶されている過去200msecの音声パラメータについての類似度も計算する。

このように入力遮断スイッチ29を構成することによつて、音声、騒音判定信号によつて音声認識の機能を停止したり機能させたりすることができるので、騒音の大きいときの誤認識を未然に防止することができる。

第9図は音声識別回路の他の例を示すブロック図であり。前述の第8図に示す音声認識回路29では、入力遮断スイッチ291を設け、騒音が大きいときにはこのスイッチ291の接点を開いて音声入力信号を音声認識回路29に与えないようにした。しかしながら、第9図に示す実施例では、入力遮断スイッチ291を設けることなく、CPU35に音声、騒音判定信号を与えるようにする。そして、騒音が大きいときにはCPU35による音声認識動作を禁止するものである。それ以外は前述の第8図と同じである。

第10図はこの発明の他の実施例の概略ブロック図である。前述の第6図に示す実施例では、前述の第(4)式の $\Phi(\tau)$ を0としたとき相互相関係数を相互相関係数計算機25で計算し、 $\Phi(0)$ の値に応じてマイクロホン17と18とに入力された音声は人間の発音した音声であるかあるいは外部からの騒音であるかを判定するようにした。しかし、この第10図に示す実施例では、第(4)式における相互相関係数 $\Phi(\tau)$ を相互相関係数計算機39で計算する。これは人間の発音した音声の場合は $\tau=0$ の近傍に $\Phi(\tau)$ が最大を有し、騒音は側方から入力されると想定しているの

14

つのマイクロホン17と18から出力される信号の遅れ時間検出器33で検出する。そして、その遅れ時間をコード化して出力し、その遅れ時間の範囲が音声であると判定すべき範囲かそれとも騒音と判定すべき範囲かを相関関数しきい値レジスタ34の値によつて比較して判定し、音声であるかを表わす判定信号を出力する。それ以外は前述の第6図と同じである。

第11図はこの発明のその他の実施例を示すブロック図である。この実施例は騒音は側方から入るとしているの

10 $\tau=0$ から離れた位置で騒音のパワーが集中することに着目して騒音の判定信号を出力するものである。すなわち、 $\Phi(\tau)$ のうち $\Phi(0)$ および $\Phi(0)$ の近傍に人間の発音が集中し、それ以外は騒音とみなすことができることに着目し、前述の第10図に示す実施例と同様に

15 して、相互相関関数計算器32によつて相互相関係数 $\Phi(\tau)$ を演算する。そして、最大値検出器35によつて相互相関係数 $\Phi(\tau)$ の最大値を求める。また、相互相関係数計算機32は相互相関係数 $\Phi(0)$ を出力し、相互相関係数 $\Phi(\tau)$ を相互相関係数(0)によつて演算器36で割算する。そして、演算器36で演算された値の大き

20 さとしきい値レジスタ37に記憶されている値と比較器26で比較し、音声であるかあるいは騒音であるかを判定するようにしたものである。それ以外は前述の第10図に示す実施例と同じである。

第12図はこの発明のさらにその他の実施例を示すブロック図であり、第13図は重み関数と音声/騒音比との関係の一例を示す図である。

構成において、マイクロホン17と18とから出力される音声信号に基づいて相互相関係数計算機25によつて相互相関係数 $\Phi(0)$ を求める部

35 分は前述の第6図に示す実施例と同じである。そして、相互相関係数計算機25によつて求められた相互相関係数は音声/騒音比対応信号として類似度計算機45に与えられる。

一方、指向性マイクロホン19から出力された音声信号はフィルタ28を介して音声パラメータ抽出器47に与えられる。なお、フィルタ28は前述の第8図に示すプリエンファシス回路292を含んでいるものとする。音声パラメータ抽出器47は同じく第8図に示すバンドパスフィルタ3

15

01ないし308、整流回路311ないし318およびローパスフィルタ321ないし328など含む。そして、音声パラメータ抽出器47は各バンドパスフィルタ301ないし308のそれぞれ周波数帯域に応じた信号の強さを出力し、それを音声パラメータメモリ38に与える。この音声パラメータメモリ38は前述の第8図に示したものと同一ものが用いられる。但し、入力部384には音声の有無にかかわらず類似度計算機45が入力部384に記憶されたデータに従って計算しているとき以外は常時音声パラメータ抽出器47から出力される音声パラメータを順次更新するように構成される。

また、前記音声パラメータ抽出器47は指向性マイクロホン19に入力された音声の音の大きさ(音圧)に対応した信号を出力し、これを音声期間判定器48に与える。音声期間判定器48はその信号に基づいて音声期間を決定し、その結果を音声期間レジスタ49に記憶する。そして、この音声期間レジスタ49の内容は類似度計算機45に与えられる。類似度計算機45は音声期間レジスタ49に記憶されている音声期間に基づいて、音声の始端からたとえば200msec以前と音声の終ったことを検知し、終端を決定した時点から音声パラメータメモリ38の入力部384に記憶されていてかつ音声期間レジスタ49によつて区間が設定された音声パラメータのデータと、音声パラメータメモリ38に登録されているキーワードの音声パラメータのデータとの間の類似度の計算を開始する。この計算において、類似度計算機45は相互相関係数計算機25から出力される音声/騒音比対応信号を重み関数として類似度の計算を行なう。

なお、この実施例では、音声パラメータ抽出器47から得られる音声パラメータは、5 msecごとに8種類のデータを取込み、たとえば1秒あたり200回の時間軸に沿ったデータとして得るようにしている。ここで、計算量を減少するために、時間の変化の少ない部分のデータを捨てることによつて情報量を一定量まで減少させるような処理を行なつておく。今、1回の取込むデータをN種、時間軸方向のデータの長さをM個とし、音声パラメータメモリ38の登録部383に記憶されている音声パラメータの要素をTij、入力された

16

音声の音声パラメータの要素をPijとしたとき、よく知られたユークリッド距離dは

$$d = \sum_{j=1}^M \sum_{i=1}^N (T_{ij} - P_{ij})^2 \quad \dots (14)$$

として表わされる。ここでiは1回に取込まれたN種うちの種類に、jは時間軸方向のデータの順序にそれぞれ対応している。通常、類似度と距離(この場合にはd)とは、逆数の関係がある。この実施例では、音声/騒音の比率に対応したW(SN)なる重み付けを行なつたユークリッド距離dWは、

$$dW = \sum_{j=1}^M W(SN) \cdot \sum_{i=1}^N (T_{ij} - P_{ij})^2 \quad \dots (15)$$

を用いて類似度を求めている。重み関数W(SN)は一例として、第13図のようなものが用いられている。但し、第13図において横軸はSN比、縦軸は重みを表わしている。

このようにして得られた類似度から、あまりに類似度が小さい場合には判定不能であるとするが、ある値以上の類似度がある場合には最大値の類似度を有するキーワードであるとキーワード判定器46において判定し、インタフェイス31にそのキーワードに対応した信号を送出し、ラジオ7の制御を行なう。なお、類似度計算機45およびキーワード判定器46は、たとえばマイクロコンピュータなどによつて構成できる。

上述のごとく音声/騒音比によつて重みを変えて距離または類似度計算を行なうことによつて、SN比が大きいときの音声パラメータ同士の比較を優先的に行ない、SN比が小さいときの値はなるべく無視するように構成することによつて、音声/騒音比のよくないときの認識性能を向上できかつ瞬時的に音声/騒音比が低下したときの誤認識を特に減少することができる。もちろん、第13図に示した横軸のS/Nは前述の第5A図で示したS/Nを用いてもよいことは言うまでもない。また、第13図において重み関数W(SN)は連続的に変化する関数としたが、簡単なためには、たとえばS/N=17dBから上を1、下を0とするような2値による重み付けを行なつてもよい。

第14図はこの発明のさらにその他の実施例を示すブロック図である。この第14図に示す実施

17

例は、前述の第12図に示す類似度計算機45に与える音声/騒音比対応信号に代えて前述の第10図に示す音声、騒音判定信号を与えるようにしたものである。そして、類似度計算機45は音声、騒音判定信号が2値であるため、重み関数W(SN)としても2値として出力する。

第15図はこの発明のさらにその他の実施例を示すブロック図である。この第15図に示す実施例は、前述の第12図に示す類似度計算機45に与える音声/騒音比対応信号に代えて前述の第10図に示す演算器43からの信号を用いるようにしたものである。そして、割算器43からの信号に基づいて、重み関数の値を変えて距離または類似度計算を行なう。もちろん2値関数として扱つてもよいことは言うまでもない。

なお、前述の第10図、第11図、第14図および第15図に示す相互相関関数計算機39として、前述の第(3)式および第(4)式における周期Tを無限大とする極限を求めないで、ある有限区間において数値計算する方法がよく用いられる。この場合、データの長さはマイクロホン17と18との間隔(30cm程度)の空気の伝播時間の2倍(約2 msec)より長ければ計算できるが、精度を上げるために10倍以上長くするように構成される。

なお、上述の実施例において、マイクロホン17と18は運転者11の口許から等距離にあるものとして説明したが、これは必ずしも必要でなく、2つのマイクロホン17と18から出力される運転者の音声信号の時間差をずらせて前述の説明と同様に行なうことができることは言うまでもない。

また、上述の説明では、この発明を車のラジオ7を制御するための音声識別装置として説明したが、これに限ることなくその他の制御対象を制御するために用いてもよいことは言うまでもない。

18

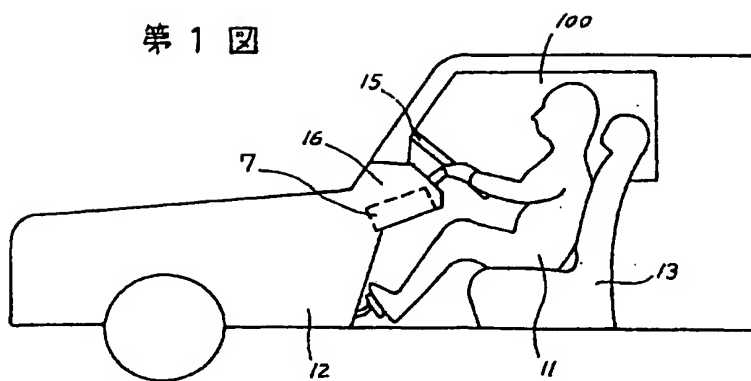
以上のように、この発明によれば、比較的高い騒音の中においても人間の発音した極めて正確に識別することができる。

図面の簡単な説明

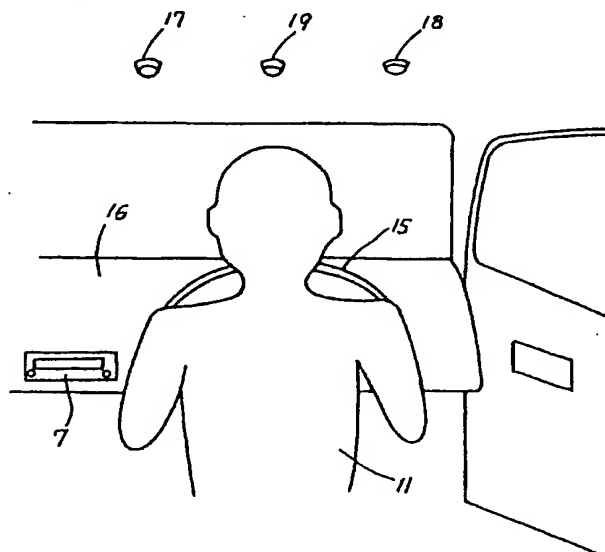
第1図はこの発明の一実施例が適用された自動車を横方向から見た概略図である。第2図は同じく運転者の後方向から見た概略図である。第3図は同じく自動車の車内の概略図である。第4図は音声信号と騒音との関係を説明するための図である。第5A図および第5B図は相関係数の計算結果の例を示す図である。第6図はこの発明の一実施例の概略ブロック図である。第7A図および第7B図は相関関数計算機の一例を示す概略ブロック図である。第8図は第6図に含まれる音声認識回路の具体的なブロック図である。第9図は音声認識回路の他の例を示すブロック図である。第10図はこの発明の他の実施例の概略ブロック図である。第11図はこの発明のその他の実施例を示す概略ブロック図である。第12図はこの発明のさらにその他の実施例を示す概略ブロック図である。第13図は重み関数と音声/騒音比との関係の一例を示す図である。第14図はこの発明のその他の実施例を示すブロック図である。第15図はこの発明の他の実施例のブロック図である。

図において、17、18、19はマイクロホン、25は相互相関係数計算機、26は比較器、27、41、44はレジスタ、29は音声認識回路、31はインタフェイス、38は音声パラメータメモリ、39は相互相関関数計算機、40は遅れ時間検出器、42は最大値検出器、43は割算器、45は類似度計算機、46はキーワード判定器、47は音声パラメータ抽出器、48は音声期間判定器、49は音声期間レジスタを示す。

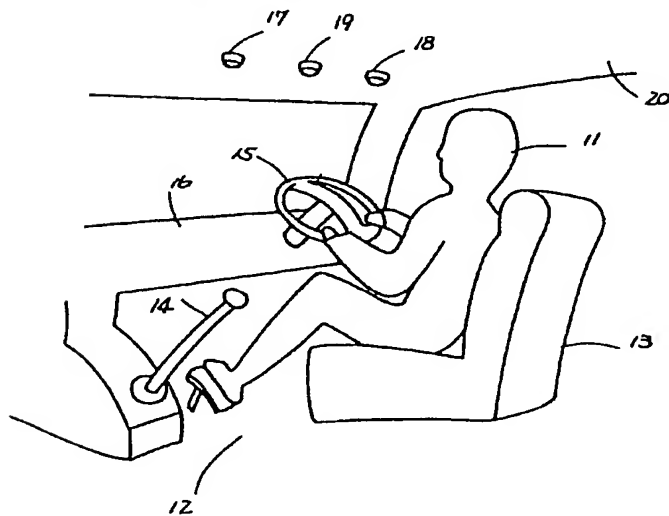
第 1 図



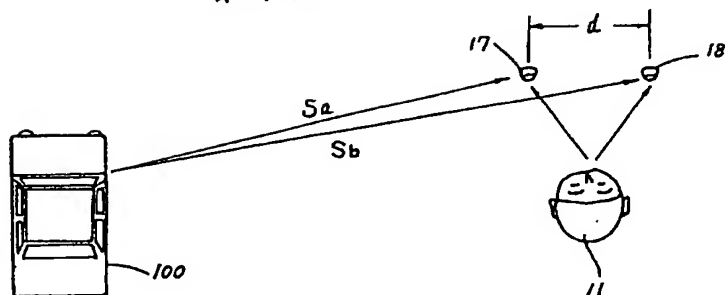
第 2 図



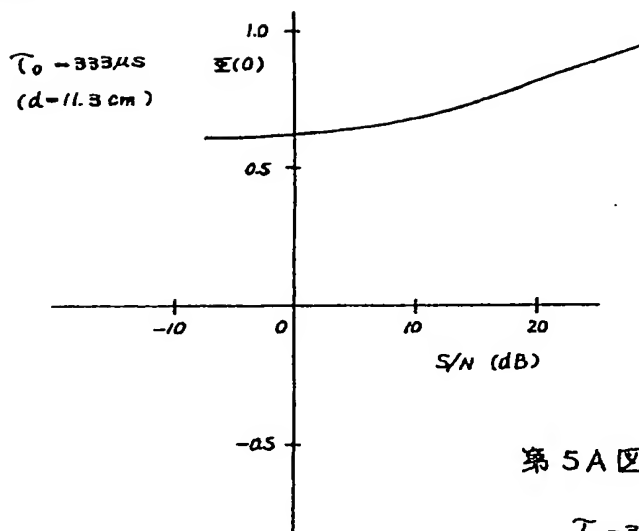
第 3 図



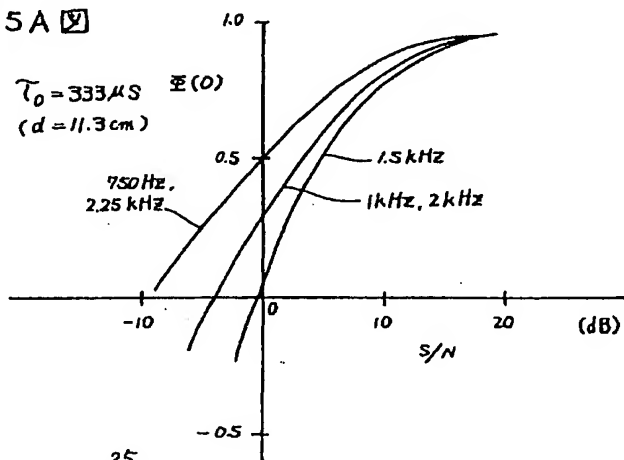
第 4 图



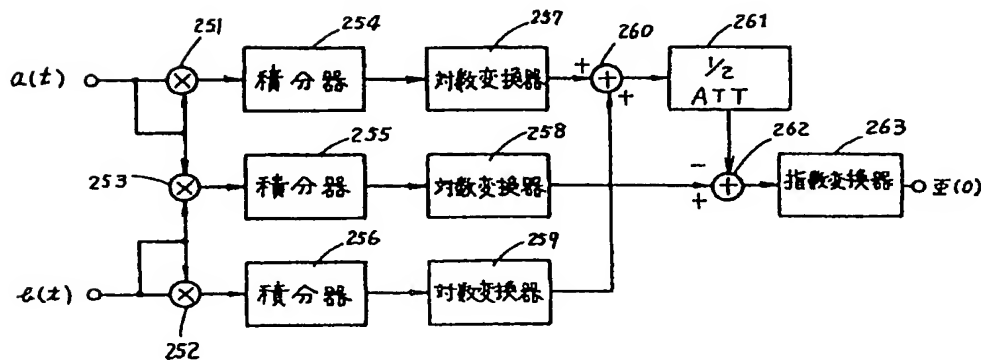
第 5 B 图



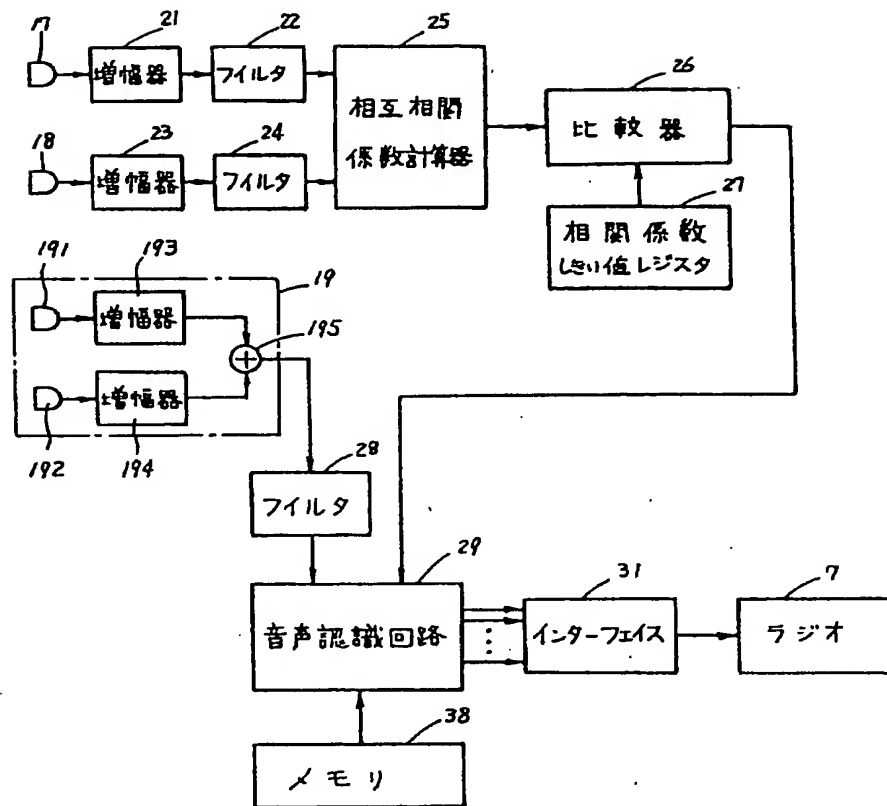
第 5 A 图



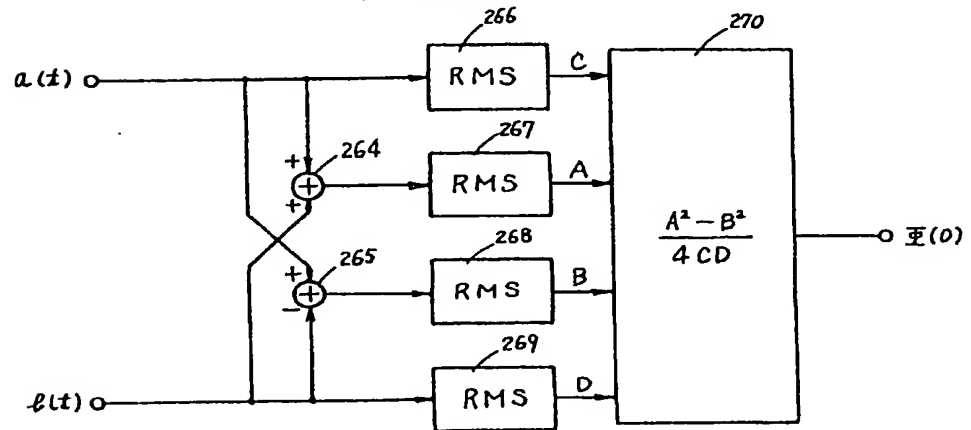
第 7 A 图



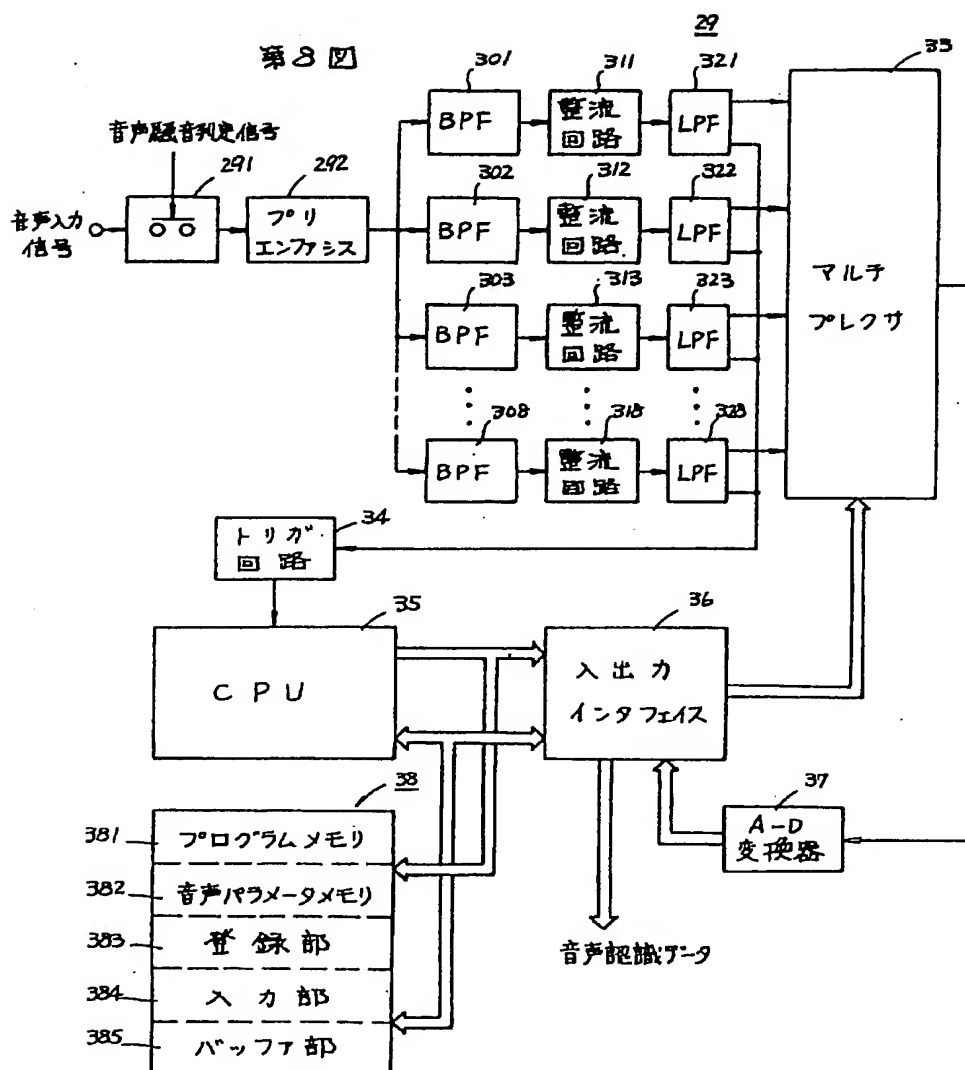
第 6 図



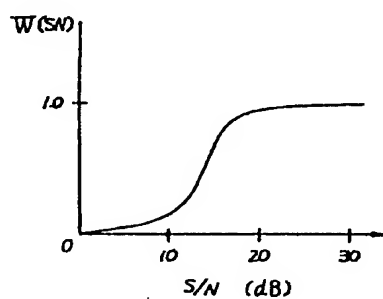
第 7B 図



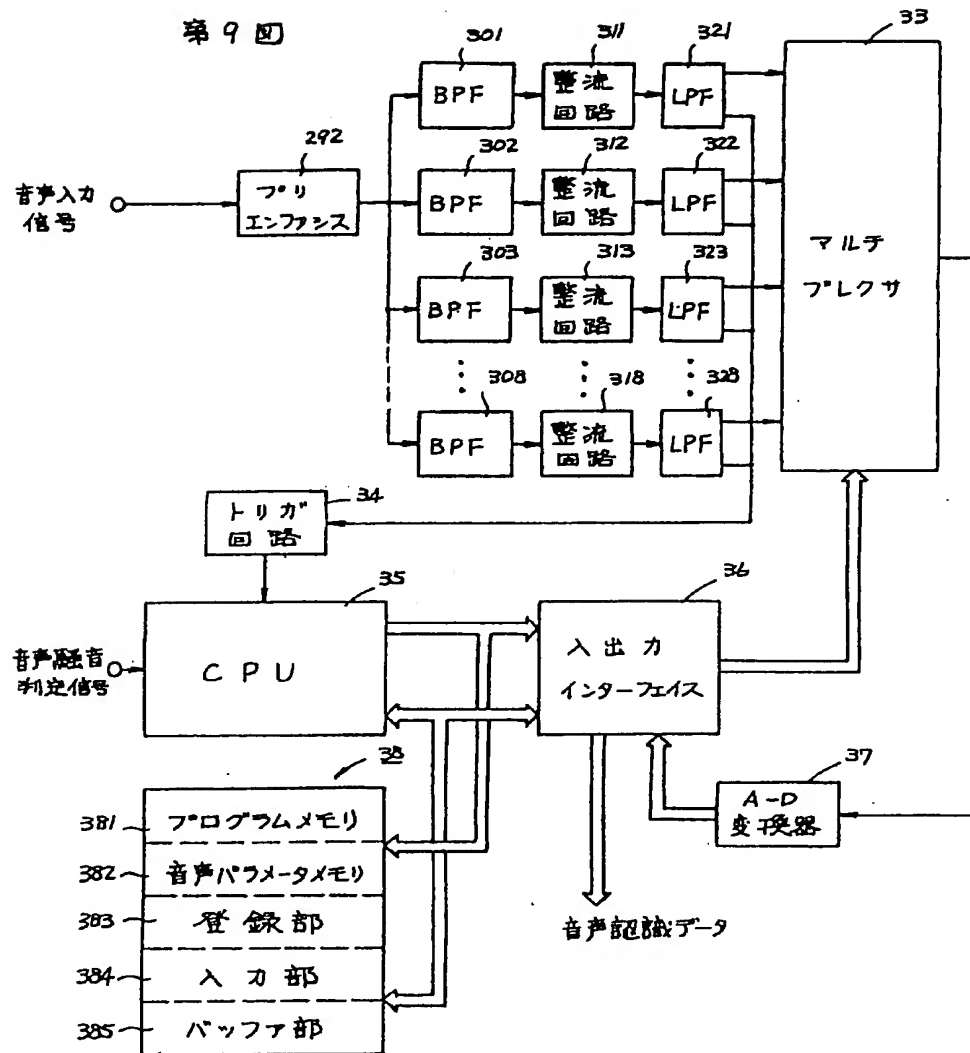
第8図



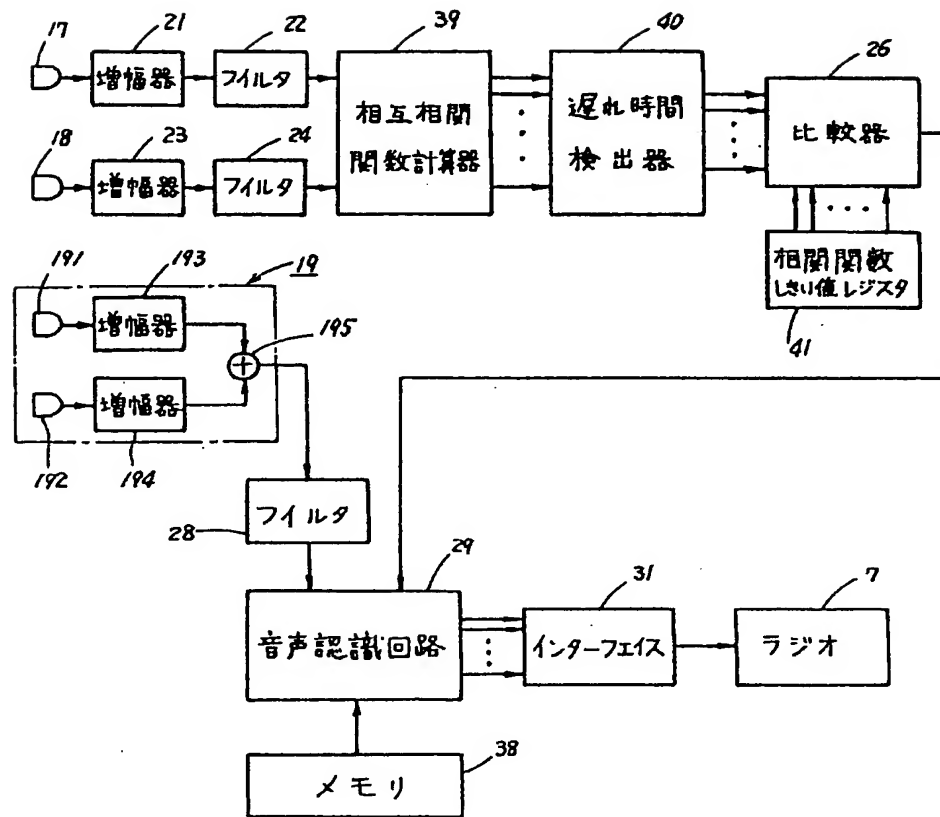
第13図



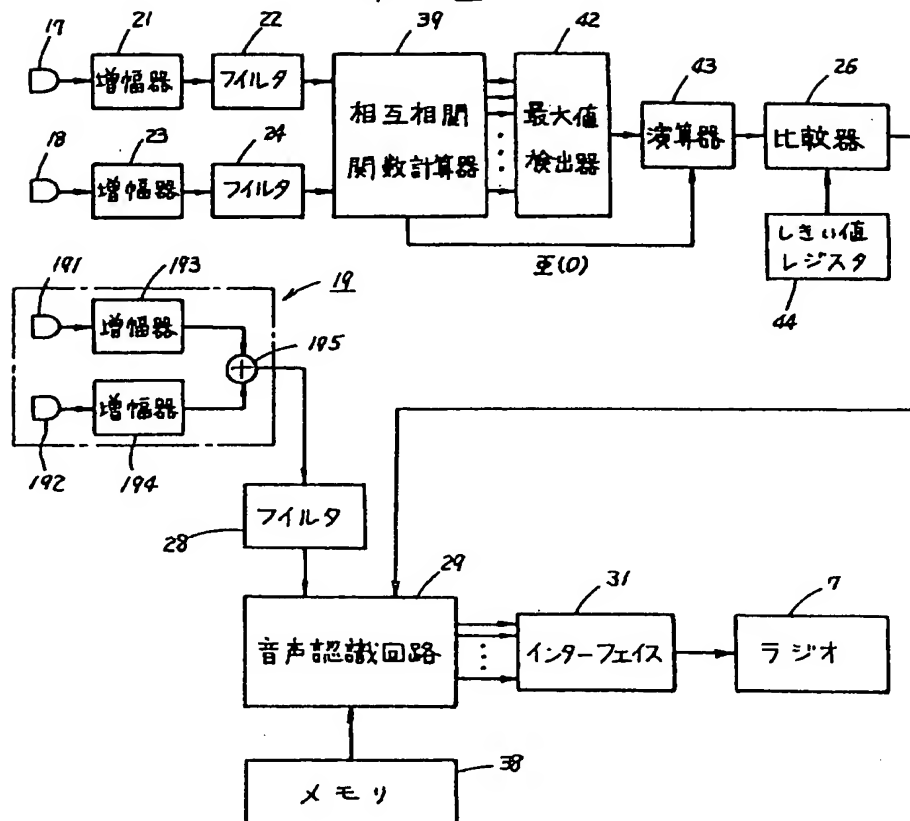
第 9 図



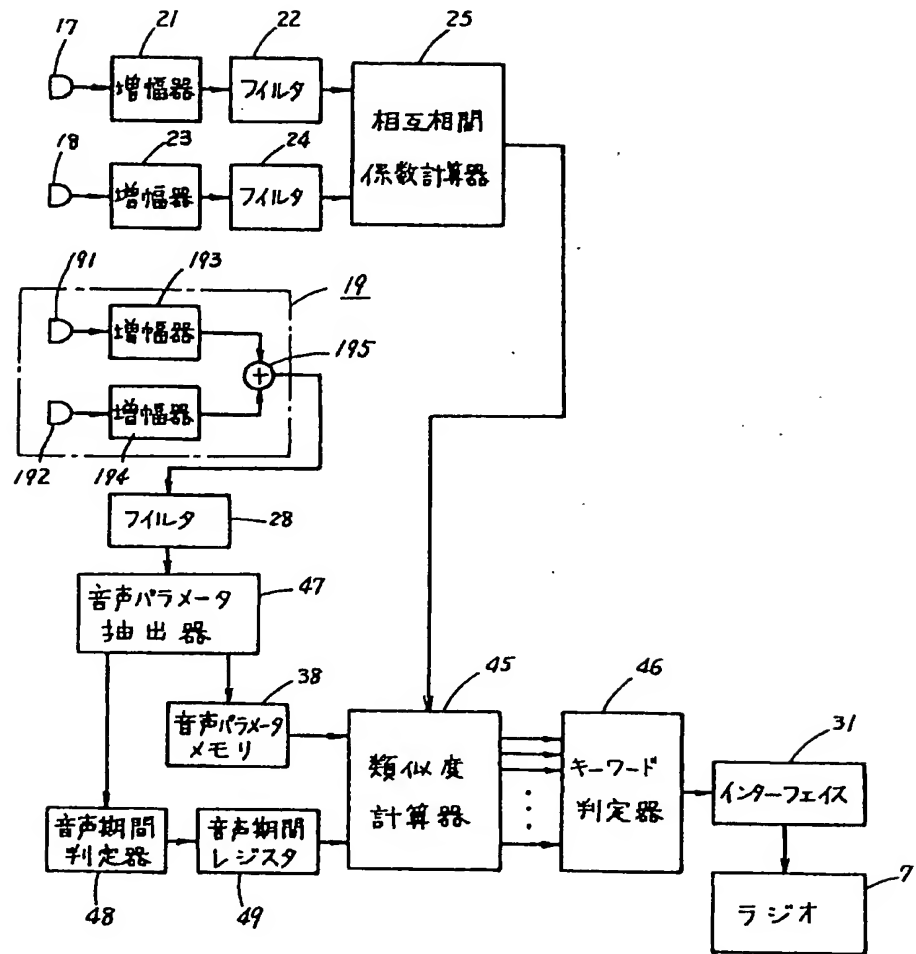
第 10 図



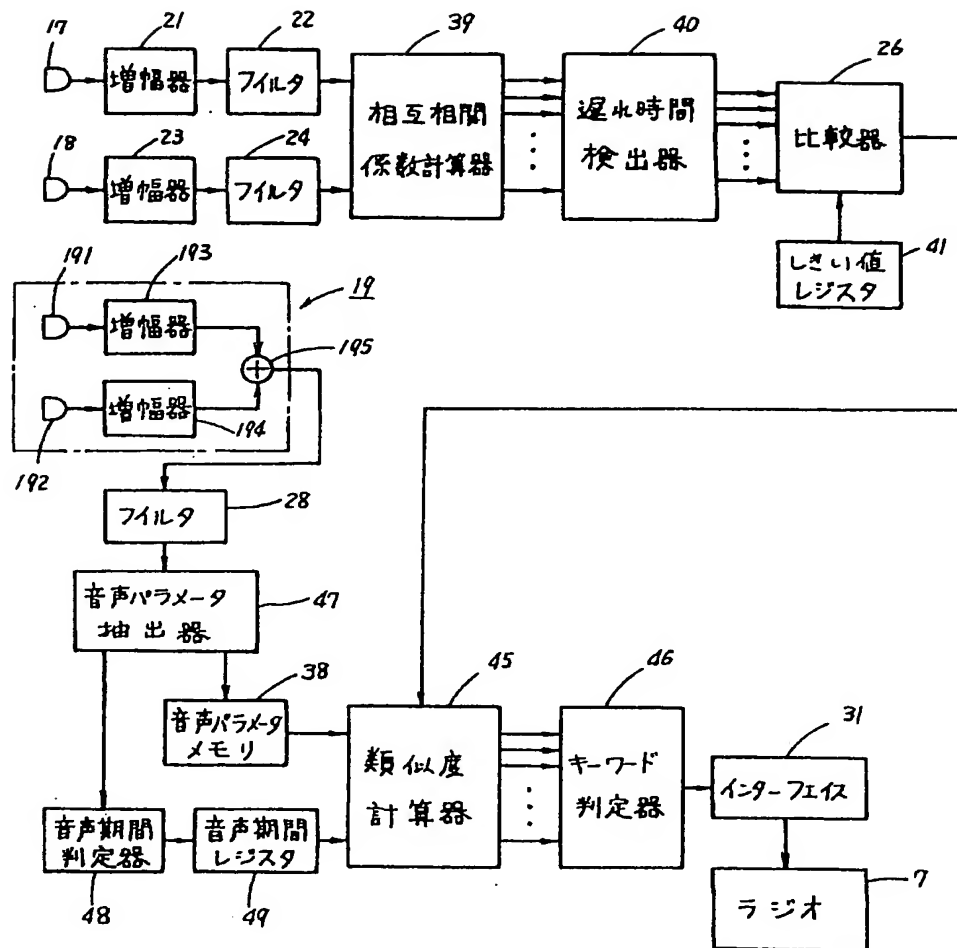
第 11 図



第12図



第 14 図



第 15 図

